EXHIBIT A

TARGET TOPIC

AUTOMOTIVE

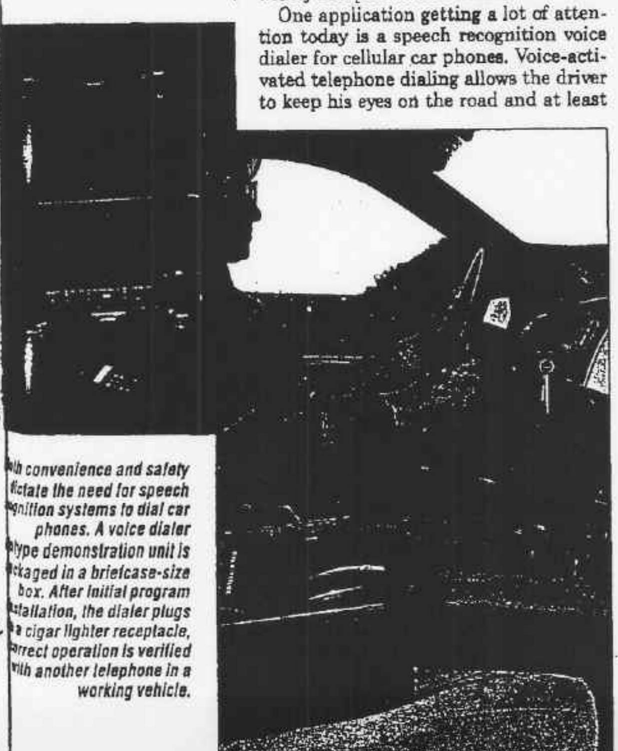
DIALING A PHONE BY VOICE

SAVARAJ I. PAWATE
PETER EHLIG
Technical Staff Members
Texas Instruments Inc.
Houston, TX

Soon you may be able to "dial" a car phone and turn on the lights and wipers with voice commands.

Look for speech recognition to be the next hot technology in the burgeoning automotive electronics industry. In fact, some experts expect voice command systems that control vehicle functions to become widely accepted in this decade.

One application getting a lot of attention today is a speech recognition voice dialer for cellular car phones. Voice-activated telephone dialing allows the driver to keep his eyes on the road and at least one hand on the wheel. Conventional dialers, in contrast, require operators to look at a keypad to punch in numbers, a dangerous activity in moving vehicles.



With convenience and safety dictate the need for speech recognition systems to dial car phones. A voice dialer prototype demonstration unit is packaged in a briefcase-size box. After initial program installation, the dialer plugs into a cigar lighter receptacle. Correct operation is verified with another telephone in a working vehicle.

The voice dialer recognizes both male and female voices, as well as a number of dialects. It can have a vocabulary of 25 or more words, depending on memory size. Surprisingly, all this functionality requires only one digital signal processor (DSP).

The voice dialer employs a speech recognition algorithm known as continuous density Hidden Markov Modeling (HMM). HMMs are statistical models for vocabulary words. The algorithms devised to decode voice patterns require substantially more computing power than other techniques, but the improved recognition accuracy outweighs any added expense incurred by using bigger microprocessors.

The voice recognition system has a speaker-independent mode, which means a person does not have to train it to learn his or her voice. For example, any rental-car customer can use the dialer. Any American speakers, regardless of their accents, can be accommodated. Continuous speech recognition is employed so the speaker can talk naturally; no deliberate pauses between words are required.

In addition to unsurpassed accuracy, the voice dialer solves a related communications problem. The cellular telephone industry is rapidly running out of available channels because of the demand for such service. However, a new algorithm called Vector Sum Excited Linear Predictive (VSELP) speech coding, allows the

TARGET TOPIC

phone system to accommodate more channels in the available bandwidth than previous methods.

Using the dialer

A typical application uses a grammar definition program built into, or down-loaded to, the DSP memory, so either man or woman can speak to a car tel phone and say "Call office" or "Ca home." He or she can also state the nu ber to be called, using the words zer through nine for digits or the word "o for zero. The user can also define a repe tory name, for example, "Call Harvey."

The heart of the dialer comprises fixe point DSPs, a ROM-based design partic larly suited for cellular phones. The DS has a number of built-in hardware fea tures that speed the implementation speech recognition algorithms. Con sequently, the phones make full use state-of-the-art digital technology t maximize available telephone chann bandwidth.

Voice dialing features can be added t cellular telephones by simply increasi system memory — other DSP devices a not required. The single speech codin DSP can be time shared to handle voi recognition as well because both fun tions do not need to run simultaneous Further, integrated cellular telephone can use the same DSP to control othe functions, such as vehicle entertainme equipment, climate, and windshiel wipers.

Voice dialer RCM and RAM com binations can be varied to handle diffe ent size boot programs, program memor and data. The programs differ dependin on the number of telephony applicatio and the functions provided. An anal interface to the telephone handset, a alpha-numeric display, and interrupt driven connections to the telepho handset complete the set-up.

New product development

To aid in the design of new speech re ognition products, the dialer doubles as development system. An RS-232 inte face, for example, supports download external software and provides a cond for control and input information other systems associated with the dial As a result, the voice dialer is easily in grated into a specific application en ronment or another development syst and evaluated.

The RS-232 port downloads to a se rate 64k RAM in the voice dialer. The b program transfers the downloaded p gram and data to the correct DSP memo

The dialer has uses other than the phone application. They include perso computers or workstations where vo

EVERYTHING OLD IS NEW AGAIN

Speech recognition technology is not new. A speaker verification system for military security was introduced in 1974, several years after research began in the 1960s. Even then, the system was said to be superior to fingerprint identification. TI also used a version of the system to control entry to its own computer center.

Today, speech and development systems are designed for a variety of applications, including text-to-speech, record/playback, telephone management, language recognition and speaker verification. Also, credit card verification systems are now widely used.

Text-to-speech algorithms convert ASCII text (as it appears on a computer monitor) into spoken English. The computer-generated voice is natural, intelligible, and has an unlimited vocabulary. Specific applications include inventory assessment, order entry input, and status review.

Record/playback applications are similar to tape recorders or dictation machines. The user can record notes, speeches and other material. However, computer storage provides greater clarity than magnetic recordings and enables the recorded file to be easily merged with other data files.

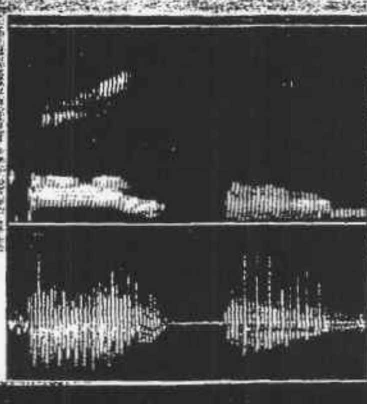
Telephone management systems employ computers to answer telephones, replay messages, and dial other telephones. Applications can be more complex than simple voice mail. In computer banking, for example, customer transactions phoned in can be confirmed at each step of the process by a synthesized voice.

Language recognition enables a computer to recognize complete sentences as they are spoken. One system, for example, can handle applications requiring up to 2,000 words. Language recognition goes beyond mere word recognition; entire sentences are analyzed using context analysis to help determine what is spoken. Language recognition is particularly useful in applications where keyboards cannot be used.

Speaker verification identifies a person through his or her unique voice characteristics. As such, it is ideal for a wide variety of security or entry control applications.

Successful applications arise from a melding of research and development in speech and semiconductor technology, and speech algorithms. For example, established multiple speech databases help create speaker-independent models for the digits used in the voice dialer.

Speech application development requires special software and hardware tools and utilities, and run-time libraries. Such software is available for a variety of DOS and Unix platforms. For example, a speech system tool kit (Speech System V) is available for Xenix or Unix systems running on Intel 80386-based computers. The tool kit also contains an interface for Unix systems operating on minicomputers.



DSP algorithms recognize the digitized form of an analog speech pattern. The top waveshape is a speechgraph of the words call home. The lower waveform is a spectrogram of the same phrase.

...y, so either
...k to a car tele
...ffice" or "Cal
...state the num
the words zer
the word "oh"
...define a reper
...all Harvey."
...omprises fixed
design particu
...hones. The DSP
hardware fea
...lementation o
...rithms. Con
...ake full use o
technology to
...phone channe
...an be added to
...mply increasing
DSP devices are
speech coding
...handle voice
...se both func
simultaneously
...ular telephones
trol othe
...rtainment
...nd windshield
...nd RAM com
...handle differ
...ogram memory
differ dependin
...ny application
...ded. An analo
...ne handset, an
and interrupt
the telephon
up.

...elopment

...ew speech rec
...ler doubles as
...n RS-232 inter
...ts downloaded
...vides a condu
information to
with the dialer
...r is easily inte
...pplication env
...opment system
...loads to a sep
...aler. The boo
...oaded pr
...DSP memor
...er than the c
include person
...ons where voic

recognition is used instead of keyboard input. Also, voice input can supplement factory automation and process inspection data for various machines and computers.

A speech recognition system can also provide hands-off control of a vehicle entertainment system, climate control, windows, windshield wipers, and door locks. For example, a driver can select a radio station with his voice or change the interior temperature without removing his hands from the steering wheel. The voice system can also query the vehicle for fuel status and mpg ratings. Even more elegant features can be had at negligible cost, such as a voice lock that allows the vehicle to be started only by authorized persons.

A demonstration voice dialer system is contained in a portable, briefcase-size box. It is powered by either a 220/110-Vac supply or 12 Vdc through a vehicle cigar lighter receptacle. Such a portable voice dialer can be used as a development system or a test set to diagnose faults in telephones in other mobile units.

The voice dialer circuit is located on one printed-circuit board with programmable array logic (PAL) to minimize the number of individual support logic chips. Voice dialer subsystems include analog circuits and codec, processor and RAM memory, processor control and EEPROMs, display and communications port, and power.

Application-specific grammar

An algorithm can be loaded that makes the dialer recognize up to 25 words without discriminating male or female voices. And application-specific grammar can be either downloaded to the system through an RS-232 serial port, or installed at the factory.

A grammar is also called a sentence model. The DSP and speech recognition algorithms understand and respond to sentence models, and control the syntax by which the words are put together.

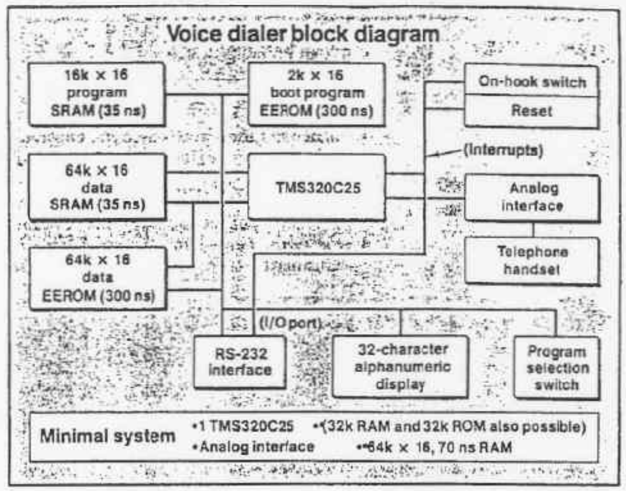
After the grammar is loaded, the voice dialer recognizes the following sequence of commands spoken in any order; call office, call home, or number (digits).

In this sequence, number is a digit string of any length, for example, number 216-666-7777 is a legal sentence. A 1-s pause (or other adjustable value) terminates any speech. When the voice dialer recognizes a complete phrase followed by the pause, it displays a period (.) on the voice dialer 32-character alpha-numeric liquid-crystal display screen. The commands 'enter' or 'cancel' can also terminate the connection.

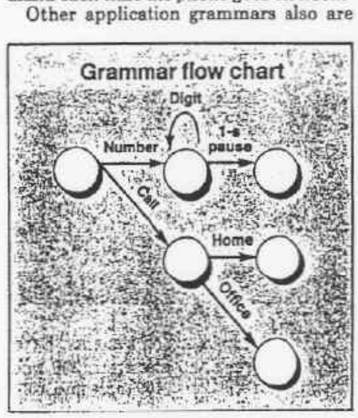
Pressing the off-hook switch on either the voice dialer case or the handset restarts the voice recognition process. In fact, the system recognizes just one command each time the phone goes off hook.

Other application grammars also are possible. An application may, for example, require that the speech recognition system recognize names and the word call as in the command call John Jones.

A basic voice dialer vocabulary consists of 11 digits (zero through nine and the word oh for zero) and four words (call, office, home, and number). But other words are easily added to the application grammar. In one version of the dialer,



The voice dialer requires either a TMS320C25 or TMS320C51 DSP with data memory, program memory, and EEPROM. A telephone handset interface, RS-232 port, display, and various switches comprise a system with a digital configuration that is different for each speech recognition algorithm that it employs.



Flow chart shows operation of the voice dialer when application-specific grammar is loaded. Here, the commands call office, call home, and number (digits) are possible, where digits is a digit string of any length.

AUTOMOTIVE

TARGET TOPIC

other common words used are enter, cancel, area, code, extension, and emergency.

The database connection

Many speaker-independent word models were created for the voice dialer to eliminate a training phase needed by earlier speech recognition systems. By collecting speech samples from 200 native American speakers (100 male and 100 female), statistical models for each vocabulary word were created. Thus, the likelihood of an unrecognizable word was largely diminished. Care was taken to sample different geographical regions to reflect various dialects. The repertory of voice information is archived in a speech database.

Recognizing that different accents need to be accommodated in certain applications, a speaker-adaptive operating mode was developed. In this mode, the voice dialer boots up with a speak er-independent model. The model is "seed" and the voice dialer controlling al gorithm continuously adapts the mod to the user in what is called a voice diale training mode.

Many novel applications also are possi ble using the database concept. For exam ple, a vocabulary may be developed that specific to one automobile manufactur or customer. For some applications, su as a personalized car phone that is d abled when others try to use it, TI c supply speaker-dependent capability f a code word.

In the present voice dialer, all need voice recognition functions, such as HM algorithms, signal processing, and gra mar control are performed by one DS For more complex applications, howeve such as large vocabularies and more co plex grammars, more than one DSP m be needed. Multiprocessor architectu allows algorithm partitioning so th larger vocabularies may be recogni and accommodated.

Experimental versions of a mul processor DSP architecture for speech r ognition have already been made. many as 32 DSPs were connected which present, uses an IBM AT computer a host for development and input/outp functions. ■

DSP TARGETED FOR SPEECH RECOGNITION

The newest DSP, the TMS320C51, has an architecture especially configured for speech processing. The design speeds speech algorithm processing much as a hardware multiplier/accumulator speeds more conventional DSP signal processing.

For example, an important calculation performed by a speech algorithm is selecting a maximum or minimum value out of a set of values. Recognizing this, designers implemented such a maximum and minimum instruction set in hardware for the TMS320C51. A description of the maximum value instruction that compares only two values is illustrated here to help understand the more complex operation for a set of values.

Assume that the maximum value of two numbers is to be found. One is placed in the TMS320C51 accumulator, the other is placed in the accumulator buffer. The instruction CRGT initiates the following sequence: The contents of the accumulator are compared to the contents of the accumulator buffer, and the larger (signed) value is loaded into both registers.

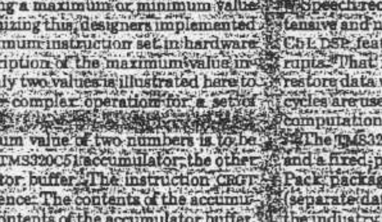
A carry bit is modified according to the comparison result. For example, if the contents of the accumulator are greater than or equal to the contents of the accumulator buffer, the carry bit is set to 1; otherwise it is zero. A similar procedure finds the least value in a set of values using the CRLT instruction.

A hardware feature of the TMS320C51 that makes it particularly suited to voice recognition is that, unlike other DSPs, the C51 performs single-cycle 16 × 16-bit multiplications in 35 to 50 ns. Data shifting and address manipulations also are in hardware rather than microcode or software.

Speech-recognition algorithms typically are arithmetic intensive and need to access as much DSP power as possible. The C51 DSP features a zero-overhead context switch on interrupts. That means no extra cycle time is needed to save or restore data when an interrupt is received. Because no timing cycles are used for data save/restore, that time is available for computation.

The TMS320C51 is a fifth-generation digital signal processor and a fixed-point machine. Available in a 132-pin Quad-Flat Pack package, the 5-V static CMOS Harvard architecture (separate data and program buses) DSP can be tested using the industry standard JTAG IEEE P1149.1 boundary scan logic.

Capable of more than 28 Mips, the DSP features on-chip ROM, program/data RAM, dual-access data RAM, and memory security. Also on-chip are address-mapped software wait state generators, serial ports, a hardware timer, five internal and four external user-maskable interrupts, and 64k I/O ports accessed by sixteen 16-bit address lines.



Texas Instruments Speech System V Toolkit is a software development package used with a 80386-based computer to create speech programs. The tool kit provides the environment to make systems for voice recognition, record-and-play, text-to-speech, and telephone management. An option is also available for speaker verification applications in security products.